12

Language Documentation & Conservation Special Publication No. 15
Reflections on Language Documentation 20 Years after Himmelmann 1998
ed. by Bradley McDonnell, Andrea L. Berez-Kroeker & Gary Holton, pp. 122–131
http://nflrc.hawaii.edu/ldc/
http://hdl.handle.net/10125/24814

Reflections on diversity linguistics: Language inventories and atlases

Sebastian Drude
University of Iceland
Goethe-Universität Frankfurt

This contribution gives a short overview of "language inventorying": research aiming at creating comprehensive catalogues and atlases of all the languages in the world, which has seen a boost with the renewed interest in linguistic diversity triggered by the awareness of language endangerment in the 1990s. By focusing on the development of the ISO standard 639 and SIL's Ethnologue, the main advances and issues in this area are discussed. The overview concludes by presenting the major alternative resources, in particular Glottolog.

The label "diversity linguistics" has been introduced by Martin Haspelmath and others at the Max-Planck-Institute for Evolutionary Anthropology in Leipzig.[1] To my knowledge, it was first used in the context of the final conference of that institute's department of linguistics (MPI-EVA 2015). Now there exist a number of activities under this label, including a book series "Studies in Diversity Linguistics" (Haspelmath 2014ff). In a broad sense, the term designates those branches of linguistics that show interest in the diversity of languages, their structure and relationship: descriptive linguistics (especially of previously understudied languages, often in a fieldwork setting), language typology, and comparative linguistics. Language Documentation is included in or at least a close neighbor to this group.

In a narrower sense, the term could refer to those studies that are interested in the diversity in itself, aiming in a first step at creating comprehensive catalogues of the world's linguistic diversity, where all languages are recorded with the information necessary to identify them, and with additional information. We label these studies here "language inventorying", and this is the focus of this contribution.[2]

[1]I would like to thank my colleagues in the ISO working group TC37/SC2/WG1 on language coding, and some colleagues who provided me with information and assessments in personal comments, in particular Harald Hammarström, Chris Moseley and Martin Haspelmath.

[2]Sometimes "Language taxonomy" is used, but this would imply genetic classification, which is often, but not necessarily, part of language inventorying.

ISBN: 978-0-9973295-3-7

Among the most basic identificatory pieces of information needed for each language in a language inventory are its name(s) and its geographical distribution, which is why language maps/atlases together with catalogues are the major tools and products of language inventorying. The additional information to be provided for each language is in principle an open-ended set of properties. Common are the genealogical classification and relationships with other languages, the number of speakers and other aspects of the 'language context', including their linguistic vitality, how well studied and described they are, and how much documentation (in the modern sense of annotated multimedia collections of language use, cf. Himmelmann 1998; Gippert, Himmelmann & Mosel 2006) is available. It can also include further structural / typological properties.

The tradition of language inventorying goes back to the Renaissance with Gesner's (1555) "Mithridates" (referring to the famously polyglot antique Persian/Greek king Mithridates VI). With the increasing outreach of the European empires, this work has been continued in the late 18th and early 19th centuries, for instance with the "Catálogo de Lenguas…" by the Jesuit Hervás y Panduro (1800 ff). The perhaps best known work of this kind was published by Friedrich Adelung and his successor Johann Severin Vater (1806 ff); in reverence to Gesner, it was also called "Mithridates", and contained a detailed presentation of almost 500 languages with text samples (mostly the Christian "Our Father" prayer). These works informed the formation of what later would become the discipline of linguistics, in personalities such as Wilhelm von Humboldt, one of the most distinguished early scholars interested in diversity linguistics.

The 20th century has seen a few scholars interested in language inventorying, beginning with "les langues du monde" by A. Meillet and M. Cohen (1952 and earlier editions back to the 1920s), and perhaps most notably Voegelin & Voegelin (1977). Also, several original regional overview works were compiled, such as Wurm and Hattori (1981) on the Pacific, or Sebeok (1977) on the Americas. Other compilations covered only a selection of major languages, such as Ruhlen (1987). A milestone was the comprehensive atlas of languages edited by Moseley and Asher (1994).

The interest in having a clear picture of the world's linguistic diversity, and thus in language inventorying, arose again in the 1990s, when the linguistic community at large became aware that this very diversity is vanishing at a worrying pace (best known are the lead articles in Language 1992 by Hale and others; most notably Krauss 1992). The same trigger also led to the establishment of language documentation and to the revival of interest in diversity linguistics, after decades of focusing on linguistic theories, the most prominent of which usually abstracted away from inner and outer linguistic diversity, deeming it enough to study English and perhaps a few other languages.

In the 1990s, the most comprehensive and best known catalogue of the world's languages was the "Ethnologue", published by the Summer Institute of Linguistics (now 'SIL International'). The Ethnologue started out in the 1950s (first edition 1951), compiled first by R.S. Pittman, as a checklist, so to speak, for the Wycliffe Bible Translators (WBT), the sister organization and main sponsor of SIL International that aims at translating the bible in all languages of the world. Maps were included in the fourth edition (1953), and the register grew from a few dozen languages and language groups to several thousand, mostly relying on information provided by SIL/ WBT missionary-linguists. B. Grimes took over as general editor around 1970 and transformed it into a general reference work (now published by SIL), systematically filling gaps, including major languages and those where no first-hand knowledge by SIL/ WBT members was available, also by simply including other work such as Voegelin & Voegelin (1977) or Wurm & Hattori (1981). One major

enhancement was the introduction of unique three-letter codes in the 10th edition (1984), allowing to unequivocally refer to languages despite the notorious confusion involving language names. Its 1996 edition, the oldest one which still can be obtained online (Grimes 1996), Ethnologue listed 6703 living languages,[3] a number which increased comparatively less since then. When Grimes handed the editorship over in the early 2000s, it was around 6800; the latest five (now annual) editions list just around 7100 languages.

The need for referring unambiguously to languages arose together with the new interest in diversity linguistics: now also linguists were in need of 'checklists' for statistics on language diversity, on language endangerment, on coverage of description—and of documentation. This need increased with the rapid technological developments: for instance, software companies need to offer localization of their products in ever more languages; international bodies such as the WWW consortium and UNICODE need to refer to individual languages. Also, with the rising language documentation efforts, language archives and similar institutions need to identify the materials they host, for instance for search engines and portals such as the Open Language Archive Community (2003ff) to be able to aggregate information.

An international standard for this purpose is provided by the International Standardization Organization ISO under the number 639. In fact, this is a group of standards. Its first part, now named ISO 639-1, was established in 1967, it contains (now ca. 200) two-letter codes like "en" for English. The second part ISO 639-2 was approved in 1998; it contains three-letter codes for now around 400 major individual languages[4] and some 70 codes for groups of languages (mostly either genealogical, e.g. "afa" for Afro-Asiatic languages, or geographical, e.g. "cau" for Caucasian languages). This part 2 came from two sources (terminologists and librarians) which were harmonized.[5]

With the new needs to unambiguously refer to all the other languages in the world, ISO approached SIL international in the early 2000s to include the Ethnologue's three-letter codes as ISO 639 part 3, and to serve as the "registration authority" that maintains and updates this part of the ISO standard. In its 15th edition (2005), organized by R. G. Gordon, Jr., the Ethnologue had adopted the ISO 639-2 codes for all languages (replacing hundreds of conflicting codes), and in 2007, ISO 639-3 was established based the Ethnologue codes. Since then, Ethnologue officially follows ISO in the question what counts as a language.

With this move, the ISO 639 standards came under the attention of diversity linguists at large, and received much critique of different kinds (most prominently perhaps by Morey, Post & Friedman 2013; see also the reply Haspelmath 2013). It is worthwhile to discuss some of these issues in detail as several of them are also relevant for other language inventorying efforts.

[3]There are contradictory numbers. The online version speaks of 6703 languages; Wikipedia (WikiProject Languages 2015) counts 6883 primary language names, and the list of language codes and names distributed together with the 13th edition in 1998 has 7825 lines/entries (Grimes 1998).

[4]'Major language' may be interpreted as "ausbau language" (language by development) in the sense of Kloss (1967): varieties that developed a literary standard and serve official functions, among others. (These are among the criteria to receive a code in ISO 639-1 and 639-2.) In this conception, most other languages are "abstand languages" (languages by distance)—a linguistic variety or group of varieties characterized of being not mutually intelligible with any other variety or language. We have no space here to argue that the concept of "language" designates real entities which are more than an ideological construct, even though the identification of individual languages may be tricky, for instance in the case of dialect chains, see below.

[5]In some 20 cases, two synonym codes were admitted. This happened where the terminologists were using codes mostly based on the autonym of the languages (e.g. "eus" for Basque, or "fra" for French) while the librarian's MARC codes were based on the English name ("baq" for Basque, "fre" for French).

Certain issues concern the institutional setting and use. For instance, government or funding and other agencies can mistake the fact of a language to be listed or not as being an "authoritative" statement of its status or even existence. This is a problem for any such standard and any language inventory, whoever compiled it, once it is widely accepted. Its close connection with the Ethnologue and thus, a missionary organization with potentially several agendas, some open, some less so, is seen as critical by many academic linguists, especially in regions where SIL has contributed little to the local academic development but more to the weakening of indigenous cultures, and, consequently, languages through supporting fundamentalist Christian proselytizing (see the contributions by Dobrin et al. 2009). Yet, alternatives to the current setting would require a solid institutional framework, as the revision process is expensive and high technical reliability is needed for ISO 639 to be acceptable.

Other critiques concern the mnemonic character of the three-letter codes which often look like acronyms. This becomes problematic when the language name that can be identified as the base for the code is deemed inappropriate. This is true, but no good solution seems feasible at this point. More than two thirds of the 17,576 possible combinations are taken; they cannot be recycled for reasons of consistency. If all criticized labels were to be exchanged, there is a real risk of running out of codes (new codes hardly get any mnemonic match anyways). ISO certainly will (and arguably can) not get into the merits of appropriate labels, as many political issues are involved—for instance, who is authorized to complain, and who to decide? The downside is that codes may be rejected by language communities. The perhaps most prominent case is the Mapudungun language, many of whose speakers reject the code "arn" which reminds of the earlier name Araucanian, considered offensive. This, however, makes creating a Wikipedia in Mapudungun problematic, as the use of the ISO code is required.

More fundamental critiques concern the delimitation of languages versus dialects or other varieties, or language families, and the heterogeneous criteria (linguistic and socio-political) applied, even more so as this is a dynamic field where languages die and new varieties and eventually languages emerge. The former points question the very feasibility of language inventorying as such. The latter point (dynamics) is not a serious objection if everybody is aware that any such list is bound to regular revision.

It is true that not only languages as defined as comprising all varieties that are (possibly serially) mutually intelligible (abstand languages in Kloss' sense) are listed in ISO 639-3 (and the Ethnologue), and even not only abstand languages and ausbau languages (see footnote 4), but also many dialects. For instance, there are 18 High German (without Yiddish) and 11 Low German (without Dutch and other Low Franconian varieties) entries although most linguists would agree that there are at most two (ausbau) languages (High and Low German) in one single German dialect continuum.[6] Also in the case of small language communities, different mutually intelligible dialects have repeatedly received separate entries and codes if they are spoken by ethnically / politically separate groups, as they function as an emblem of ethnic identity.

Still, to give up on creating a comprehensive language inventory on such grounds would mean to throw the baby out with the bath water. In fact, even in (at the first glance) complex and confusing situations involving, in particular, dialect continua, the specialists who work with those varieties usually can come to an agreement if they agree on the

[6]One can argue that some varieties spoken by former German emigrants/colonists (such as pdc, Pennsylvania German, or hrx, Hunsrik) or other geographically separated communities (e.g. cim, Cimbrian) have developed into separate abstand languages.

criteria (linguistic, mainly mutual intelligibility, and socio-political). There will always be borderline cases, but the general situation can in principle be improved and eventually solved by (a) making both the linguistic and socio-political criteria for each entry explicit and (b) improving the criteria and conceptual framework, recognizing that the concrete language topology may be more complex than can be captured by the simple conceptual triad "language family – language – dialect". The only recent consistent attempt known to me to advance at the second front is offered by T. Kaufman with his classifications of South American languages (Kaufman 1990; 2007), going back to earlier proposals by Hockett (1958). In addition to the most common and paradigmatic case of *families / languages / dialects* (some languages being *dialect chains*), he distinguishes *language areas* and their constituent *emergent languages* (between them there are clear boundaries but high intelligibility), and *language complexes* (these are dialect chains with *virtual languages*: subsets functioning as languages). There may be other elaborate conceptions that I am not aware of, but none of these seems to have influenced current works of language cataloging, where the trichotomy of dialect / language / language family are the only concepts used, usually not even sufficiently distinguishing between languages on linguistic vs. languages on socio-political grounds.

In my view, the most interesting questions concern ISO's (and the Ethnologue's) factual adequacy. A very detailed assessment with hundreds of individual corrections has been offered by Hammarström (2015). With regard to scientific standards, the most pertinent critique against the Ethnologue is that most of its data is not verifiable, as sources often remain unknown. The second complex concerns the genealogical classification. (This is not a problem of ISO 639-3, which does not engage in classification at all.) Ethnologue has created genealogical trees which are frequently at odds with the best researched historical-comparative work. In Hammarström's (2015:734) evaluation, the expert-like-ness is only around 30%. In this context, the "authoritative" character of the Ethnologue is indeed potentially dangerous as readers often will not take the trouble to look for expert classifications, and even scientific publications such as the International Encyclopedia of Linguistics (Frawley 2003) just follow the Ethnologue without verifying the merit of individual genealogical trees.

As to the question of identifying and counting languages, central to this essay, the Ethnologue has often been criticized to overcount (mainly by recognizing too many dialects as languages; see the example of German varieties above). Dixon (2012:463–4) even estimates that the number of languages if consistently grouping all mutually intelligible varieties together in one language, one would arrive at a count of 'a good deal less' than 4000 languages. Hammarström (2015:733) again provides the arguably best informed estimate of 6497 known languages (solely on the grounds of mutual intelligibility, but dividing larger dialect continua in a number of different languages, see Hammarström (2005)), or between 5,593 and 7,400 languages with a confidence interval of 95%. Of these, around 6000 are living. Conclusively, by purely linguistic criteria, the Ethnologue overcounts indeed, but only by some 10–15%.

Still, there are other problems with the Ethnologue and ISO 639-3. The very relationship between the two is problematic, because when one presents the Ethnologue with shortcomings and mistakes (many are individually listed and substantiated by Hammarström), Ethnologue can point to ISO 639-3 as the authority on which languages to list; Ethnologue is supposedly only following. ISO 639-3, in turn, will point to its change request submission mechanism, which puts the burden of doing the paperwork for correcting each of the mistakes on the specialists or language inventoryists like

Hammarström, and the outcome is uncertain (and the referees or experts that make the decisions for SIL as registration authority are not known)[7]—unless you work closely with the registrar of ISO 639-3 (or the editors of Ethnologue?) to get many changes accepted without having to publish the evidence. In addition, the Ethnologue now is behind a paywall, which makes much of its information, in particular the language maps, much less accessible (see the discussion by Skirgård 2016). (ISO 639-3 remains openly accessible.)

In view of these limitations, what are the alternatives, in particular from the academic community? Luckily, there are several, albeit each with its scope and problems. Limitations of space prevent me to discuss each of the existing relevant resources that have been created over the last twenty years or so, largely as a result of the renewed focus on diversity linguistics and fieldwork and language documentation efforts.

It is clear that such resources should not be created by one person or small group alone, but needs contributions from many experts. This is well demonstrated by the *linguasphere registry* (Dalby, Barrett & Mann 1999), which shows a resourceful conception, many valid insights and much knowledge, but is preliminary for many individual languages / families / areas (and does not cite any sources) (Vajda 2001). Its 8-level language codes have not been widely taken up, and it has been rejected as the basis for a planned part 6 of ISO 639 on language varieties.

There are strong arguments that suggest that the most advanced and sound catalog of the world's languages is the Glottolog (Hammarström et al. 2018; see also Hammarström 2016). It lists all languages with their genealogical and geographical position according to the best information available to the editors.[8] More importantly, it lists many important sources for each language. On the other hand, at this point Glottolog gives no language context information such as speaker numbers or vitality, but they provide a reliable assessment of the degree of description (they plan on including more, see Hammarstrom et al. in prep.), also presented in the comprehensive "Handbook of Descriptive Language Knowledge" (Hammarström 2018).

Glottolog assigns codes (the 'Glottocode') of letters and ciphers, which are being taken up by important other resources such as Wikipedia. The authors were partly also involved in creating other resources important for diversity linguists, in particular WALS (Dryer & Haspelmath 2013). Glottolog is currently actively maintained, but does not seem to have a sustainable long-term institutional basis; as happens with most such resources: if the current editors decide to stop maintaining it, there is no guarantee that it will be kept alive. Nevertheless, at this point it is arguably the most complete and reliable language catalogue, surpassing other works in particular in its genealogical classification and balanced judgment as to identify languages versus dialects, but it needs to be complemented with other sources for language context information.

Two further major resources are also well-researched and do provide language context information, but are restricted to endangered languages. The UNESCO Atlas of the World's Languages in Danger (Moseley 2010) was first edited as a book (Wurm 1996), has been largely enhanced in the early 2000s and is still updated by its main editor, although the funding currently barely allows to keep the atlas accessible.Chris Moseley is also co-editor of the Routledge Atlas of the World's languages (Moseley & Asher 2007), itself

[7]Besides this problem, the transparency has increased a lot since 2007; for a few years now all decisions are publicly visible. In the 2018 round of change requests, anyone can see and comment the requests, which improves the possibility for experts and community members to interfere considerably.

[8]Based on the geographical information (unfortunately, just a coordinate, the geographical 'centre-point') it is possible to create a world map of the languages according to Glottolog, see Caines et.al. (2016; realized as Caines & Bentz).

arguably one of the most complete language catalogues, although its reliability varies among different areas.

UNESCO is planning on a new, much more comprehensive and ambitious edition including all languages. The UNESCO atlas applies its own endangerment measurement based on its experts panel's conception (UNESCO Ad hoc Expert Group on Endangered Languages 2003), which is different from (and not rarely comes to different results than) Ethnologue's EGIDS scale (Lewis & Simons 2010).

The Endangered Languages Project (ELP) (Alliance for Linguistic Diversity 2013ff) has set up a website where users can contribute with information and language resources such as recordings. The underlying technology has been provided by Google.org, the research by NSF grants led by the LinguistList and the University of Hawaii. Underlying the ELP is ELCat, the Endangered Languages Catalog, developed mainly at the University of Hawaii. It has, again, its own scale of language endangerment which also indicates the certainty of the assessment, an element that is desirable for all such assessments.

The LinguistList itself holds a number of relevant resources, in particular MultiTree (LinguistList 2014), which collects genealogical trees and thus provides genealogical and further information on many languages and dialects, and LL-Map (LinguistList, n.d.), a similar collection of published language maps which links individual languages to the MultiTree resource.

This overview can only offer a selection of the most prominent work done in the field of language inventorying. More and more initiatives emerge, such as perhaps most recently the Wikitongues project (Wikitongues), similar in spirit to the Endangered Languages Project, more inclusive (not only endangered languages) but with less ties to the academic community.

Overall, we have seen much progress over the last few decades in our efforts to catalogue and map the languages of the world. We now can say with much confidence that there are around 6000 living languages in world if we apply mainly the linguistic criterion in a way so that varieties between which no or low mutual intelligibility exists are kept apart as different languages even in a dialect chain situation; and we know of some 500 more which are not spoken any more. If we allow for socio-political and functional criteria, we can identify at least around 7500 languages, and we can list them and refer to them with unambiguous codes.

As far as we can see at this point, what is needed is: (a) a movement to unite or combine the existing language catalogs with a quality check, for instance by establishing a recognized clearing house or permanent expert panel with an accepted academically recognized procedures to evaluate the status of individual languages; (b) more reliable language maps which better represent the complex and often multilingual language landscapes than dots or mostly non-overlapping colored areas can; (c) a more sustainable long-term institutional backing and funding so that the valuable resources created are reliably available in the future.

References

Adelung, Johann Christoph & Johann Severin Vater. 1806. *Mithridates oder allgemeine Sprachenkunde – mit dem Vater Unser als Sprachprobe in bey nahe fünf hundert Sprachen u. Mundarten.* Reprint: Hildesheim: Olms, 1970. Berlin: Voss.

Alliance for Linguistic Diversity. 2013ff. Endangered Languages Project. (`http://www.endangeredlanguages.com/`) (Accessed 1 May, 2018)

Caines, Andrew & Christian Bentz. Languages of the world. (`https://cainesap.shinyapps.io/langmap/`) (Accessed 1 May, 2018)

Caines, Andrew, Christian Bentz, Dimitrios Alikaniotis, Fridah Katushemererwe & Paula Buttery. 2016. The Glottolog Data Explorer: Mapping the world's languages. Paper presented at the *VisLR II Workshop*.

Dalby, David, David Barrett & Michael Mann. 1999. *The linguasphere register of the world's languages and speech communities. 1st ed.2 vols.* Hebron, Wales, UK: Published for Observatoire Linguistique by Linguasphere Press / Gwasg y Byd Iaith.

Dixon, R. M. W. 2012. *Basic linguistic theory volume 3: Further grammatical topics.* Oxford, New York: Oxford University Press.

Dobrin, Lise M., Jeff Good, William L. Svelmoe, Courtney Handman, Patience Epps, Herb Ladley & Kenneth S. Olson. 2009. SIL International and the disciplinary culture of linguistics. (Ed.) Lise M. Dobrin. *Language* 85(3). 618–658. (`doi:10.1353/lan.0.0132`)

Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *The World Atlas of Language Structures (WALS) Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. (`http://wals.info/`) (Accessed 1 May, 2018)

Frawley, William (ed.). 2003. *International encyclopedia of linguistics, Second edition.* Oxford, New York: Oxford University Press.

Gesner, Conrad. 1555. *Mithridates. De differentiis linguarum tum veterum tum quae hodie apud diversas nationes in toto orbe terrarum in usu sunt observationes. Neudruck: Aalen: Scientia, 1974, hrsg. u. eingel. v. Manfred Peters.* Zürich: Froschoverus.

Gippert, Jost, Nikolaus P. Himmelmann & Ulrike Mosel (eds.). 2006. *Essentials of language documentation.* Berlin, New York: Mouton, De Gruyter.

Grimes, Barbara F. (ed.). 1984. *Ethnologue, 10th ed.* Dallas, Texas: Summer Institute of Linguistics.

Grimes, Barbara F. (ed.). 1996. *Ethnologue, 13th ed.* Dallas, Texas: Summer Institute of Linguistics. (`http://www.ethnologue.com/13/`) (Accessed on 30 April, 2018)

Grimes, Barbara F. 1998. Language names (by code). (`http://xml.coverpages.org/langcodesEth.txt`) (Accessed 30 April, 2018)

Hale, Kenneth, Michael Krauss, Lucille J. Watahomigie, Akira Y. Yamamoto, Colette Craig, LaVerne Masayesva Jeanne & Nora C. England. 1992. Endangered languages: On endangered languages and the safeguarding of diversity. *Language* 68. 1–42.

Hammarström, Harald. 2005. Counting languages in dialect continua using the criterion of mutual intelligibility. *Journal of Quantitative Linguistics* 15. 34–45.

Hammarström, Harald. 2015. Ethnologue 16/17/18th editions: A comprehensive review. *Language* 91(3). 723–737.

Hammarström, Harald. 2016. Linguistic diversity and language evolution. *Journal of Language Evolution* 1(1). 19–29. (`doi:10.1093/jole/lzw002`)

Hammarström, Harald. 2018. Handbook of Descriptive Language Knowledge. (https://www.academia.edu/3142333/Handbook_of_Descriptive_Language_Knowledge) (Accessed 2 May, 2018)

Hammarström, Harald, Sebastian Bank, Robert Forkel & Martin Haspelmath (eds.). 2018. Glottolog 3.2. Jena: Max Planck Institute for the Science of Human History. (http://glottolog.org/) (Accessed 1 May, 2018)

Hammarström, Harald, Robert Forkel, Thom Castermans, Bettina Speckmann, Kevin Verbeek & Michel A. Westenberg. In prep. Visualizing language endangerment and language description hand in hand.

Haspelmath, Martin. 2013. Can language identity be standardized? On Morey et al.'s critique of ISO 639-3. *Billet*. Diversity Linguistics Comment. (https://dlc.hypotheses.org/610) (Accessed 1 May, 2018)

Haspelmath, Martin (ed.). 2014ff. Studies in Diversity Linguistics (book series). Berlin: Language Science Press. (http://langsci-press.org/catalog/series/sidl) (Accessed 26 April, 2018)

Hervás y Panduro, Lorenzo. 1800. *Catalogo de las lenguas de las naciones conocidas, y numeracion, division y clase de estas segun la diversitad de sus idiomas y dialectos*. re-published as facsilime by the Biblioteca Nacional (España), Madrid, 2008.6 vols. Madrid: [s.ed.].

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36(1). 161–195.

Hockett, Charles Francis. 1958. *A course in modern linguistics*. New York: Macmillan.

ISO 639-3. ISO 639-3. (https://iso639-3.sil.org/) (Accessed 1 May, 2018)

Kaufman, Terrence. 1990. Language history in South America: What we know and how to know more. In Doris L. Payne (ed.), *Amazonian Linguistics: Studies in Lowland South American languages (Texas Linguistic Series)*, 13–73. Austin.

Kaufman, Terrence. 2007. The native languages of South America. In R. E. Asher & Christopher Moseley (eds.), *Atlas of the world's languages*, 46–76. London: Routledge.

Kloss, Heinz. 1967. Abstand Languages and Ausbau Languages. *Anthropological Linguistics* 9(7). 29–41.

Krauss, Michael. 1992. The world's languages in crisis. *Language* 68. 4–10.

Lewis, M. Paul & Gary F. Simons. 2010. Assessing endangerment: expanding Fishman's GIDS. *Revue roumaine de linguistique* 55(2). 103–120.

LinguistList. 2014. MultiTree. A digital library of language relationships. Bloomington: Indiana University. (http://multitree.org/) (Accessed 1 May, 2018)

LinguistList. n.d. LLMap: LL-MAP: Language and Location – A Map Annotation Project. (http://www.llmap.org) (Accessed 1 May, 2018)

Meillet, A. & Marcel Cohen (eds.). 1952. *Les langues du monde: Par un groupe de linguistes, 2. Ed.*. Paris.

Morey, Stephen, Mark W. Post & Victor A. Friedman. 2013. *The language codes of ISO 639: A premature, ultimately unobtainable, and possibly damaging standardization*. Handout of a talk given at the PARDISEC RRR Conference, Sidney. (https://www.academia.edu/4725536/) (Accessed 20 January, 2019)

Moseley, Christopher (ed.). 2010. *UNESCO atlas of the world's languages in danger. 3rd ed.* Paris: UNESCO. (http://www.unesco.org/languages-atlas/) (Accessed 1 May, 2018)

Moseley, Christopher & R. E. Asher (eds.). 1994. *Atlas of the world's languages*. New York: Routlegde.

Moseley, Christopher & R. E. Asher (eds.). 2007. *Atlas of the world's languages, 2nd ed.* New York: Routlegde.

MPI-EVA. 2015. Diversity Linguistics: Retrospect and Prospect | Presentations. (https://www.eva.mpg.de/linguistics/conferences/diversity-linguistics-retrospect-and-prospect/presentations.html) (Accessed 26 April, 2018)

OLAC. 2003ff. Open Language Archives Community. (http://www.language-archives.org/)

Pittman, Richard S. (ed.). 1951. *Ethnologue: languages of the world, 1st ed.* Dallas, Texas: Wycliffe Bible Translators (WBT). (http://www.ethnologue.com/13/) (Accessed 30 April, 2018)

Ruhlen, Merrit. 1987. *A guide to the world's languages.* Stanford: Stanford University Press.

Sebeok, Thomas (ed.). 1977. *Native languages of the Americas, volume 2.* New York, Boston: Springer.

Skirgård, Hedvig. 2016. Clarifying points on Ethnologue pay-wall and language codes. Humans Who Read Grammars blog. (http://humans-who-read-grammars.blogspot.com/2016/01/clarifying-points-on-ethnologue-pay.html) (Accessed 1 May, 2018)

UNESCO Ad hoc Expert Group on Endangered Languages. 2003. Language Vitality and Endangerment. UNESCO Document, presented at a meeting in Paris, 2003. (http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/pdf/Language_vitality_and_endangerment_EN.pdf) (Accessed 20 January, 2019)

Vajda, Edward J. 2001. The Linguasphere Register of the World's Languages and Speech Communities (review). *Language* 77(3). 606–608. (doi:10.1353/lan.2001.0197)

Voegelin, Charles F. & F. M. Voegelin. 1977. *Classification and index of the world's languages.* New York, Oxford, Amsterdam: Elsevier.

WikiProject Languages. 2015. Primary language names in Ethnologue 13. Wikipedia. (https://en.wikipedia.org/w/index.php?title=Wikipedia:WikiProject_Languages/Primary_language_names_in_Ethnologue_13&oldid=645050764) (Accessed 30 April, 2018)

Wikitongues. (https://wikitongues.org/) (Accessed 1 May, 2018)

Wurm, Stephen A. (ed.). 1996. *Atlas of the world's languages in danger of disappearing* (Theo Baumann, cartographer). Paris & Canberra: UNESCO.

Wurm, Stephen Adolphe & Shiro Hattori. 1981. *Language Atlas of the Pacific Area: New Guinea area, Oceania, Australia; 2, Japan area, Taiwan-Formosa, Philippines, Mainland and insular South-East Asia.* Australian academy of the humanities.

Sebastian Drude
sdrude@hi.is
orcid.org/0000-0002-2970-7996